The Folding of the Nucleosome Chain

Karsten Rippe

6.1 Introduction

The nucleosome is the fundamental unit of chromatin and has been introduced in Chapter 3. It consists of an octameric protein core composed of two copies each of histone proteins H2A, H2B, H3, and H4, and 145–147 base pairs (bp) of DNA. The DNA stably contacts the surface of the histone protein octamer core in a left-handed superhelix of almost two turns. Histone proteins have a globular part formed by three well structured α-helices and the histone tails. These are long protruding N-terminal and H2A C-terminal extensions which lack a specific secondary structure. They are subject to post-translational modifications like acetylation, methylation, and phosphorylation at numerous sites, mostly in H3 and H4 (Chapter 4).

While the core histones (or their variant forms) are stably bound to the DNA with residence times on the hour timescale, linker histone H1 and its avian counter part H5 associate more transiently with the nucleosome core particle and induce local conformational changes of the linker DNA [1–3]. As discussed in Chapters 3 and 5, the nucleosomal DNA is partly inaccessible to other protein factors, and the positioning of nucleosomes thus regulates the access to the DNA sequence. Similarly, the linker DNA between nucleosomes can be occluded by compacting the nucleosome chain, and a 50-fold difference in its accessibility to protein binding has been reported based on a comparison of dinucleosomes with a folded chain of 17 nucleosomes [4]. Thus, the organization of the nucleosome chain can regulate DNA access. The factors that determine this process as well as the resulting structures that form with chains of less than hundred nucleosomes are discussed in the present chapter. The organization of larger chromatin regions (i.e., thousands of nucleosomes and more) is covered in Chapters 9, 17, and 20.

In numerous experiments it has been demonstrated that an extended chain of nucleosomes with ~10 nm diameter can reversibly fold into fiber structures with a diameter of 20–40 nm at physiological salt concentrations [2, 3, 5]. This conformation has been termed the 30 nm chromatin fiber. Its structure remains controversially discussed, and various models for its conformation have been
proposed [2, 3, 5–8]. It is noted that the reported diameter of chromatin fibers varies between 20 and 45 nm as measured for chicken erythrocyte chromatin and for fibers reconstituted in vitro [2, 3, 8–11]. In addition, the degree of fiber compaction covers a broad range. It is usually reported as a linear mass density of the number of nucleosomes per 11 nm fiber contour length. Values of 6–7 nucleosomes for chicken chromatin [10–12] and 8 for mudpuppy erythrocyte chromatin [13] were reported, while the corresponding value in yeast was much lower at about 1–2 nucleosomes per 11 nm chain length [14]. For fibers reconstituted in vitro values of up to 10 or even 17 nucleosomes per 11 nm fiber were measured for fully compacted chains with saturating amounts of linker histones [9]. Thus, experimental determinations of very basic parameters like fiber diameter and mass density clearly indicate that different fiber conformations exist. These appear to depend on a number of parameters like the spacing of nucleosomes referred to as the nucleosome repeat length (NRL) [2, 3, 9, 15–17], the presence and type of linker histones [2, 3, 6, 9, 16], the ionic conditions [10, 12, 18–20], and the acetylation state of histone H4 residue lysine 16 [21, 22]. Furthermore, it has been questioned whether the 30 nm fiber persists within the nucleus in an environment that is highly enriched with nucleosomes [23–26]. However, a number of studies provide evidence for fiber structures in the nucleus [2, 27–30]. While this issue remains to be addressed in further experiments there is no doubt that fiber structures form with native chromatin fragments in vitro [31–34]. By dissecting this process inherent features of the nucleosome chain can be identified that determine its conformation. If these are known in sufficient detail, the organization of the chain can be derived for a particular set of conditions, including those that are relevant within the nucleus. Accordingly, the protein–protein and protein–DNA interactions that govern the folding of the nucleosome chain into various types of structures are key parameters of the conformational properties of the genome.

6.2 Experimental Systems

6.2.1 Native Chromatin Fragments

Native chromatin fiber fragments are typically isolated from cells by a partial digestion with micrococcal nuclease (MNase). The length of nucleosome chains isolated in this manner can be adjusted to enrich a certain fragment size but is in general below 100 nucleosomes. This approach has been used for studies of the salt-dependent compaction of nucleosomes, where a length-dependent fractionation was conducted to obtain samples with 2 to ~60 nucleosomes [35–40]. Chromatin samples obtained in this manner are heterogeneous with respect to the DNA sequence, post-translational histone modifications, and the presence of non-histone proteins. However, some approaches exist to study native chromatin
formed on defined DNA sequence fragments in the length range of 5–15 kb [12, 41, 42]. A frequently used chromatin source are chicken erythrocytes since purification of relatively large amounts is straightforward [10–12, 37, 40, 43–46]. Native chromatin fragments were also purified from rat liver [32, 35, 36, 46], cow thymus [38] or brain tissue [36], sea urchin sperm [47, 48], yeast [49], and immortalized mammalian cell lines like HeLa [39, 50]. As discussed in Chapter 5, the NRL varies between species and cell type from 154 and 237 bp corresponding to a DNA linker of 10–100 bp between two nucleosomes. For example, NRL values have been determined at around 154 bp in *Schizosaccharomyces pombe*, 165 bp in *Saccharomyces cerevisiae*, 175 bp in *Drosophila melanogaster* and *Caenorhabditis elegans*, 185 bp in *Homo sapiens* [2, 51, 52], 212 bp for chicken erythrocytes [2], and 237 bp in sea urchin sperm [48]. The distribution of spacer lengths is not random but follows a ~10 bp periodicity [53], which closely resembles a helical turn of DNA (10.4 bp). This points to sterical requirements of nucleosome spacing to be compatible with the higher-order folding of the nucleosome chain. Furthermore, calculations based mainly on data of recent knockout studies demonstrate a linear relationship between the ratio of H1 per nucleosome and the NRL. The results show that the presence of the linker histone leads to a lengthening of the NRL by 37 bp [6]. In mammals the typical NRL is around 200 bp but can show large variations between tissues [2]. The NRL can be determined from a partial MNase digestion and subsequent analysis of the DNA length distribution by gel electrophoresis with an accuracy of 1–2 bp [2]. These experiments suggest that the region of regular nucleosome spacing that can be identified as a set of distinct bands comprises less than ~10 nucleosomes. This is consistent with the result from genome-wide mapping of nucleosome positions, as discussed in Chapter 5.

Given the above large variations in NRL as well as other chromatin features, it appears likely that the conformational heterogeneity of the corresponding chromatin fragments is significant. For example, chicken erythrocyte chromatin with an NRL of 212 bp appears to represent a more repressive overall conformational state and is enriched in 30 nm chromatin fibers [29] with a mass density of about 6–7 nucleosomes per 11 nm fiber [10, 11]. It contains the avian specific linker histone type H5 instead of H1. Nucleosomes from chicken erythrocytes display a characteristic stem-like structure in which H5 mediates the association of the two DNA segments leaving the nucleosome core particle over a distance of 3–5 nm before the linker DNA diverges [11, 44]. It is unclear if this type or DNA organization by linker histones is also present in chromatin from other sources. In contrast, yeast chromatin has unusually short nucleosome repeat length between 154 and 165 bp. Hho1p, its functional homolog to linker histones, has a second globular domain with a winged helix-turn-helix motif but lacks the equivalent to the C-terminal domain found in other eukaryotes [54]. Yeast chromatin adopts a more decondensed conformation of the nucleosome chain with a low mass density of 1.2–2.4 nucleosomes per 11 nm fiber [14, 49]. These large variations between chromatin from different organisms need to be accounted for when attempting to describe the folding of the nucleosome chain into fibers or other higher-order structures.
6.2.2

Reconstituted Nucleosome Chains

In many instances endogenous chromatin fragments comprise a mixture of different DNA sequences and are not well defined with respect to histone modifications and the presence of histone variants, linker histones, and other chromosomal proteins. To study nucleosome chains of a defined composition a number of reconstitution approaches have been developed, as reviewed previously [3, 5, 55–58]. In these experiments a gradient of decreasing salt concentration is used to deposit histone octamers and linker histones onto the DNA [59] (Chapter 3). The corresponding in vivo process involves histone chaperones and chromatin remodelers and results in a regular spacing of nucleosomes around a certain NRL (Chapters 3, 5, and 15). In contrast, the salt gradient reconstitution method leads to randomly distributed nucleosomes with natural sequences. Therefore, repeats of high-affinity binding sites for the histone octamer like the 5S DNA repeat from Lytechinus variegatus [60] or the “601” sequence determined from an in vitro selection of random DNA sequences [61, 62] are used to obtain an equal spacing. The resulting high regularity of nucleosome positions assembled in vitro exceeds that of native chromatin [63, 64].

Reconstituted nucleosome chains can be classified according to their length into three groups:

1. A tetranucleosome structure has been determined with a relatively short 167-bp repeat length in the absence of linker histones by X-ray crystallography [65]. It provides the first high-resolution structure for interactions between two nucleosomes that drives the higher-order folding of a nucleosomes chain.
2. In a number of studies, nucleosome arrays with ~12 nucleosomes were reconstituted on repeats of high-affinity histone octamer binding sites [5, 66–71].
3. To obtain longer chains with regular nucleosome spacing the use of high-affinity histone octamer binding sites like the “601” DNA sequence is critically important. One study used a nucleosome array with 48 nucleosomes and 177 bp NRL [68]. In another series of experiments arrays with 47–80 nucleosomes and 177–237 bp NRL (in 10 bp steps) were investigated with and without linker histone [9, 16, 22].

6.3

Nucleosome–Nucleosome Interactions

The driving force for folding the nucleosome chain are nucleosome–nucleosome interactions that compensate for the unfavorable energetic terms that arise from bending/twisting of the linker DNA, the electrostatic repulsion of the linker DNA and decreased conformational entropy [15, 72–74]. At a concentration above 50 mM salt, that is, at the physiologically relevant ionic strength, this interaction becomes attractive. This has directly been observed by a variety of
methods [2, 75–77]. Salt-dependent chain folding has also been investigated experimentally for reconstituted nucleosome arrays [5, 78, 79] and native chromatin fragments isolated from cells [19, 35, 40], as well as theoretically [73, 80, 81]. In the following the characterization of nucleosome–nucleosome interactions at physiological ionic strength is discussed.

6.3.1 The Strength of Nucleosome–Nucleosome Interactions

To quantitate the strength of internucleosomal interactions experimentally, force spectroscopy experiments are particularly instructive [82–87]. In these a nucleosome chain is bound at one end to a solid support and is then extended by pulling at the other end with forces in the range of 0.1–40.0 pN. From the resulting extension a force–distance curve is obtained. This type of experiment was conducted both with native chromatin fibers derived from chicken erythrocytes [83] as well as reconstituted nucleosome arrays [82, 84–87]. The various experimental studies yielded nucleosome–nucleosome interaction energies between 3.4 and 14.0 $k_B T$ for breaking a single interaction, where $k_B$ is the Boltzmann constant and $T$ the temperature. These results were corroborated by Monte Carlo (MC) computer simulations of chromatin fiber force spectroscopy experiments [88]. On a per mole basis $k_B T$ can be converted to $RT$, with $R$ being the gas constant. At room temperature $RT$ is equal to 0.6 kcal mol$^{-1}$ or 2.5 kJ mol$^{-1}$ and reflects the thermal energy available to the system for spontaneously occurring reactions. The corresponding experimentally determined nucleosome interaction energies cover a broad range from 2 kcal mol$^{-1}$ in native chromatin fibers [83] to 8 kcal mol$^{-1}$ for reconstituted nucleosomal arrays [84]. These values compare to average affinities of about $-12$ kcal mol$^{-1}$ for the specific binding of a protein to its DNA target site, and $-4$ to $-7$ kcal mol$^{-1}$ for typical unspecific protein–DNA interactions [89]. As discussed in Ref. [88] the large range of values is likely to reflect to which degree specific features of a given nucleosome chain allow for the establishment of optimal interactions between nucleosomes, as well as the solution environment. Relevant parameters include the NRL and regularity of nucleosome spacing as well the presence of linker histones and divalent cations. Under conditions optimal for interactions between nucleosomes the attractive energies are significant and similar to that of unspecific binding of a protein to DNA. However, constraints imposed by the DNA linker and/or the local nucleosome geometry oppose such a favorable alignment. This can render the effective nucleosome–nucleosome interaction energy insufficient to establish a compact chromatin fiber-like structure. For example, it was shown that the unfavorable electrostatic repulsions and DNA bending/twisting energies of a $\sim 60$ bp long linker DNA (NRL=$207$ bp) reduced the favorable contribution of nucleosome–nucleosome interaction to an effective value of $\sim 2 k_B T$ in the absence of linker histones and divalent cations [72, 88]. This results in open structures with low fiber mass density as observed by electron microscopy images for nucleosome arrays without linker histones [16].
6.3.2 Interactions of the Globular Part of the Histone Octamer Core

The nucleosome surface is determined by the core histones and associated nucleosomal DNA as a cylindrically shaped structure of about 11 nm diameter and 5.5 nm height with a heterogeneous charge distribution (Chapter 3). Several experimental findings indicate that the stacked alignment of two nucleosomes provides the most favorable conformation at physiological salt concentrations (Figure 6.1). This is inferred from nucleosome crystal structures [65, 90], studies of nucleosome liquid crystals [91], electron microscopy observations [16, 68, 92], and the nucleosome-nucleosome distance distributions determined by atomic force microscopy [93, 94]. Changes to the surfaces that interact during stacking can modulate the folding of the nucleosome chain [95, 96]. These can either result from differences in the histone protein sequences [90], post-transcriptional histone modifications like H3K79 methylation [97] or the incorporation of histone variants [96, 98, 99]. In particular, variants of histone H2A appear to significantly change nucleosome–nucleosome interactions via the H4 tail [67, 68, 96, 100]. The canonical H2A core histone provides an acidic patch that interacts with the positively charged H4 tail in the nucleosome crystal structure [68, 100]. The H2A variant H2A.Bbd lacks three acidic amino acids in this region, and its incorporation into the nucleosome chain inhibits folding [96]. However, the H2A.Z variant has an extended acidic patch, which appears to favor nucleosome–nucleosome interactions and chain compaction [99]. In addition to the stacked orientation some favorable interaction energies may also be provided by histone octamers and nucleosomes that interact in a side by side orientation (Figure 6.1) [91, 101, 102]. This type of associations is also observed for salt-dependent fiber–fiber interactions [34, 67, 78, 79, 103–105]. Both the stacked as well as the perpendicular/side by side interactions are critically dependent on histone tails, as discussed below.

---

**Figure 6.1** Estimated orientation dependence of the nucleosome–nucleosome interaction potential. The interaction energy is shown as a function of the center to center distance \( r \) for different oriented nucleosome pairs according to a potential used in coarse-grained computer simulations (scheme adapted from Ref. [74, 209]). Interaction energies vary with distance \( r \) and have been parameterized in this example to reach about 14.6 kJ mol\(^{-1}\) (stacked nucleosomes), 0.5 kJ mol\(^{-1}\) (perpendicular) and 1.2 kJ mol\(^{-1}\) (side by side) at the optimal distance.
Contributions of Histone Tails to Nucleosome–Nucleosome Interactions

Both experimental and modeling studies indicate that histone tails are important for mediating internucleosomal interactions and the folding of the nucleosome chain [22, 106, 107]. Removal of the histone tails leads to some increase of nucleosome flexibility and affects nucleosome–nucleosome interactions as well as the binding of other proteins to the nucleosome and/or its associated DNA [108–110]. The strength of these interactions can be modulated directly by post-transcriptional histone modifications (Chapter 4). These are set or removed in a dynamic manner by specific enzymes and target certain protein domains that specifically interact with the post-translationally modified histone state. In addition, acetylation of histone lysines can have a direct effect on the stability of the nucleosome core particle and on its higher-order interactions since the positively charged lysine is neutralized in the acetylated state [111]. Interestingly, an effect on the nucleosome structure has also been reported for DNA methylation [112].

In a number of studies with reconstituted nucleosome arrays, the contribution of histone tails to the folding of the chain was assessed by either removing the tails [67, 78, 103, 113–116], substituting certain amino acids [117] or by introducing post-translational histone modifications like acetylation [21, 22, 69, 93, 104, 118] or methylation [97]. The results indicate that the positively charged tails neutralize negative phosphate charges of the DNA backbone and promote interactions between neighboring nucleosomes [67, 80, 103, 119, 120]. Both the N-terminal tails of H2B and H3 mediate internucleosomal interaction, possibly by binding in the continuing groove of the DNA superhelix formed by two stacked nucleosomes [67, 79, 100, 103, 121, 122]. As mentioned above the interaction of the H4 tail with the acidic patch on the surface of H2A is particularly important for interactions between nucleosomes. Accordingly clipping off the H4 tail or its acetylation at lysine residue 16 can strongly reduce the compaction of nucleosome arrays [21, 22, 67, 104, 116]. Thus, the contribution of the histone tails to nucleosome–nucleosome interactions is significant. In a theoretical study it was concluded that acetylation of a single H4K16 can reduce its value by almost 2 $k_B T$ [107]. From a comparison of complete nucleosomes and those with trypsinized tails interaction energies of 2 $k_B T$ [123] and 5–10 $k_B T$ [114] were derived. When considering the results from computer simulations a total tail contribution of $\sim 5$ $k_B T$ to the nucleosome–nucleosome interaction energy seems to be a reasonable estimate [107, 124].

DNA Interactions with the Histone Octamer Protein Core

As discussed in Chapter 3, histone–DNA interaction maps based on the crystal structure of the nucleosome [100, 121] or on stretching experiments [125–127] identified a $\sim 10$-bp periodicity of interactions of the DNA with the histone proteins. These were assigned to 14 main interaction sites at regions where the minor groove faces inwards. A recent molecular dynamics analysis indicated that
each of these interaction sites can be considered as comprising DNA-histone contacts separated by \(~5\) bp that alternate between each of the two individual DNA strands (Figure 6.2a) [128]. This leads to a 5 bp periodicity pattern of 28 DNA interactions with the protein core in agreement with conclusions from force

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.2}
\caption{Distribution of histone octamer–DNA interactions and DNA unwrapping profile [128]. (a) Histone–DNA interaction map determined from all-atom molecular dynamics simulations. Fluctuations are shown for a period of 20 ns. Some unsymmetry is introduced in the nucleosome structure via the different locations of the two H3 tails. (b) For calculating the solid line relative microscopic binding constants per base pair were estimated from the molecular dynamics simulations in (a). The dashed line shows a nucleosome unwrapping profile calculated for homogeneous interaction energies at all DNA positions. Unwrapping probabilities were computed for an equilibrium binding constant \(K=10^{21} \text{M}^{-1}\) and a histone octamer concentration of \(10^{-6} \text{M}\). For further details, see Ref. [128].}
\end{figure}
spectroscopy experiments [125]. The partial spontaneous DNA unwrapping of nucleosomal DNA by breaking these contacts was studied by various approaches both experimentally and theoretically (Figure 6.2b) [128–134]. These yielded life times of several seconds for the closed state that were interrupted by open periods of a few tenths of seconds, in which up to 80 bp of nucleosomal DNA were exposed.

As shown in Figure 6.2b calculations of the unwrapping probability that take into account the specific pattern of DNA–histone interactions in the nucleosome depicted in Figure 6.2a indicate that unwrapping the first 20 bp on both sides of the nucleosome dyad axis occurs with an increased probability. It is noted that the unwrapping of nucleosomal DNA is also a crucial part of the activity of a specific class of enzymes referred to as remodeling complexes (Chapter 5). To evaluate the potential impact of unwrapping of nucleosomal DNA on the folding of the nucleosome chain the following questions are discussed here: (i) how strong are DNA–histone octamer interactions; (ii) what is the extent of partial unwrapping due to thermal fluctuations; (iii) is this process different for an isolated nucleosome as compared to a folded chain?

The strength of the interaction between the DNA and histone protein can be directly investigated in force spectroscopy experiments. Nucleosomal DNA unwraps from the histone protein core upon extension at sufficiently high forces [125, 127, 128, 135–141]. For an isolated nucleosome experimental studies indicate that DNA unwrapping starts already at \( \sim 3 \) pN [140, 141]. In contrast, for a chain of 25 nucleosomes no unwrapping of nucleosomes was observed below pulling forces of 5–7 pN [82, 84, 88]. Thus, the interactions between nucleosomes within the chromatin fiber stabilize the nucleosome structure and counteract DNA unwrapping. DNA unwrapping becomes significant above 5 pN applied pulling force [142, 143]. The outer turn (67 bp, 23 nm of DNA) dissociates first and more easily than the inner DNA turn (80 bp, 27 nm) [85, 86, 128, 138–140, 144]. For unwrapping the outer turn of nucleosomal DNA energetic costs of 10 \( k_B T \) [140], 15 \( k_B T \) [129], and 20 \( k_B T \) [85] have been derived from force spectroscopy as well as competitive protein binding experiments. For further unwrapping of the inner DNA turn in the nucleosome an energy barrier exists [85, 88, 138, 140, 145]. It appears to arise to some extent from higher affinity DNA–histone interactions flanking the dyad axis [138, 140].

6.5 Architectural Chromosomal Proteins and Chromatin States

In addition to the stably bound core and variant histones, a number of more transiently bound chromosomal proteins exist that affect chromatin structure. These include linker histones as well as a number of architectural chromatin proteins that organize the conformation of the nucleosome chain and its higher-order folding. Together with specific post-translational histone and DNA modification patterns these define different functional chromatin states.
6.5.1 Linker Histones

The linker histone H1 (present in five isoforms H1.1 to H1.5), its avian erythrocyte variant H5 or the yeast Hho1p homolog can interact with an additional ~20 bp of DNA flanking the nucleosome and organize the linker DNA [54, 146–148]. H1/H5 consist of three protein domains: a compact globular domain that is flanked by two highly positively charged N- and C-terminal domains [149–153]. In the free protein these are mostly unstructured and neutralize negative charges of the DNA phosphate backbone upon binding to the nucleosome to form the chromatosome [71, 154]. Linker histones elicit complex effects on chromatin conformation that have been discussed in detail elsewhere [6, 32, 151, 152, 155]. A high-resolution structure of the chromatosome is missing, but a number of models for the interaction of the linker histone and the nucleosome have been proposed [72, 150, 156–160]. From these and the available experimental data the effects of linker histones on chromatin organization appear to be related to two effects: it was shown that binding of the linker histone constrains the entry/exit angle of the DNA at the nucleosome, and induce a more compact chromatin conformation, in which gene expression is changed [11, 39, 44, 64, 151, 156, 161–166]. In addition, it was demonstrated that the NRL increases with the linker histone stoichiometry [6].

6.5.2 Other Architectural Chromosomal Proteins

The folding of the nucleosome chain is controlled by a number of non-histone proteins that are able both to compact and to open up chromatin and are thus referred to as architectural chromosomal proteins [56, 167]. A comprehensive discussion of structural changes on chromatin conformation induced by these is beyond the scope of this chapter, and only some of the most prominent factors are mentioned here:

1. Heterochromatin protein 1 (HP1) is involved in establishing and maintaining repressive state of pericentric heterochromatin [168–171]. Chromatin binding of HP1 is tightly connected to the methylation status of histone H3 at lysine 9. The N-terminal chromodomain of HP1 interacts preferably with H3 histone tails that carry a K9me2/3 modification [172–174].

2. Proteins with the high mobility group (HMG) motif can counteract linker histone-mediated chromatin compaction [175, 176]. They are classified into the HMGA [177], HMGB [178], and HMGN [179] groups that all bind to chromatin but display additional diverse functions. These include modulating the level of post-translational histone modifications (HMGN, HMGB) as well as regulating nucleosome positioning (HMGB1, HMGN1, HMGN2) [176].

3. MeCP2 is a member of the MBD protein family that, except for MBD3, bind methylated CpG sites via their MBD domain (Chapter 2) [180]. Structural studies with in vitro assembled chromatin demonstrate that MeCP2 binds to nucleosomes and compacts the nucleosome chain [180–182]. The chromatin
compaction activity requires the N-terminus of the protein and occurs independently of the DNA methylation state.

4. CTCF is a transcription factor that has an enhancer-blocking activity on certain promoters. Its chromatin organizing activities are described in a number of recent reviews [183–185]. It has been proposed that chromatin-bound CTCF acts as an insulator element between chromatin domains of different types by promoting the formation of DNA loops that constrain interactions between promoter and enhancer elements and that it positions adjacent nucleosomes.

5. Cohesin and condensin protein complexes have originally been identified as factors that compact the DNA in the mitotic chromosome, which is discussed in Chapter 18. A number of reports have revealed their additional functions as organizers of the higher order interphase chromatin structure by promoting interactions between chromatin loci in cis and in trans as reviewed recently [186, 187].

The above proteins and linker histones exert their effects on chromatin architecture in a complex network that involves direct protein–protein binding and the readout of DNA methylation and histone modifications as mentioned above for HP1 and MeCP2. Interestingly, HP1 [188, 189] and HMGB1 [176] interact with linker histone H1, while HMG proteins compete with H1 for nucleosome binding sites [175, 176]. CTCF mediates recruitment of cohesin to most of its binding sites, and the two factors have been proposed to stabilize interphase chromatins loops [183–187]. In addition, as discussed in several recent reviews, HP1 [171], H1 [190], and HMG [176] proteins themselves are subject to post-translational modifications that modulate their activity. For example it has been shown, that the linker histone H1.4 isoform binding to HP1 was enhanced by methylation of lysine 26 and reduced by phosphorylation of serine 27 [188, 189].

6.5.3 Chromatin States

Historically, chromatin has been globally classified into more compact biologically inactive heterochromatin as compared to transcriptionally active euchromatin that can be identified from the chromatin distribution on microscopy images [170, 191] (Chapter 17). Since transitions between these two states have been observed the term facultative heterochromatin was introduced. The classical example for this is the inactivation of one X-chromosome in female mammalian cells that adopts a distinct conformation state termed the Barr body, while genes on the other X-chromosome are active. Other functionally distinct regions include chromatin at the centromeres [192–194], pericentromeric heterochromatin [169, 191, 195], chromatin at the nuclear lamina (Chapter 8), telomeric chromatin [196, 197], and nucleolar chromatin (Chapter 12).

While the above type of classification is useful and reflects functionally different forms of chromatin, it lacks a systematic approach to describe the underlying molecular differences. This shortcoming has been addressed recently by several
studies that use either the protein composition or the histone modification pattern to identify distinct chromatin states [198–200]:

1. From genome-wide binding maps of 53 chromosomal proteins five major chromatin types with >100 kb in length were defined in *Drosophila* cells [198]. These include a repressive chromatin type that lacks the previously used heterochromatin markers. Furthermore, transcriptionally active euchromatin could be divided into two groups that differ in their protein composition pattern and the H3K36 methylation state.

2. Another study in *Drosophila* focused on histone modifications and categorized 18 different histone acetylation or methylation marks into nine patterns. These characterized the chromatin state with respect to chromosomes, genes, regulatory elements, and other functional units [199].

3. Two histone acetylation marks, six histone methylation modifications, and binding of CTCF were evaluated in different human cell types to identify patterns that characterize promoters and enhancers as well as their cell-type activity for gene expression [200].

It is anticipated that these and further studies can be integrated into a systematic chromatin classification that includes DNA sequence features, DNA methylation, histone modification patterns, nucleosome positions, occupancy with non-histone chromatin proteins, and gene-expression data. These could then be analyzed in terms of specific conformations adopted by the nucleosome chain and its higher-order folding. At low resolution this is already possible. For example, chromatin states defined by histone modifications can be compared with respect to their sensitivity towards DNase I as an indicator of an open chromatin conformation with reduced nucleosome occupancy [199].

### 6.6 Chromatin Fiber Conformations

The structure of the 30 nm fiber remains controversially discussed, and various models for the fiber geometry are currently under investigation. It is also noted that an alternative model to that of a continuous fiber is the “superbead” model, in which eight nucleosomes (chicken and rat liver) to 48 nucleosomes (sea urchin sperm) associate into a globular particle [201]. Furthermore, studies of fiber structures *in vitro* are typically conducted with short fragments (<100 nucleosomes) at nucleosome concentrations in the range of 1 μM [9, 68]. As discussed in Chapter 7, the nucleus represents an environment that is highly enriched with nucleosomes and DNA. Nucleosome concentrations during the interphase of the cell cycle are estimated to vary between 60 and 450 μM during interphase and can reach ~1.2 mM in the mitotic chromosome [202]. Under these conditions the 30-nm fiber conformation might resolve into a “sea of nucleosomes” [23–26]. Alternatively, more irregular and aggregate-like structures could form where nucleosomes from distant parts or from other chromosomes would intermingle. These
nucleosome–nucleosome interactions in cis and in trans become more favorable as the nucleosome concentration is raised and the chain length is increased, which facilitates its back-folding. Thus, further investigations are required to evaluate if the particular conditions present in the cell nucleus in terms of nucleosome concentration, fiber length, and solvent environment favor the organization of the nucleosome chain into alternative structures.

Various fiber conformations have been proposed based on *in vitro* studies with native chromatin fragments, reconstituted nucleosome chains as well as from studies of chromatin in cells [2, 9, 27, 29–34, 68, 203–207]. These can be classified according to their nucleosome stacking using the nomenclature of Depken and Schiessel [208]. The fiber conformation is described by two parameters as \((N_{\text{stack}}, N_{\text{step}})\): the number of nucleosome stacks is given by \(N_{\text{stack}}\), while \(N_{\text{step}}\) refers to the step size between connected nucleosome stacks. This is shown in Figure 6.3 for the example of a seven-start helix. In addition, fiber conformations may be different with respect to the orientation nucleosomes to the helix axis (tilt angle), the position of the linker histone, the degree of linker DNA bending and their ability to accommodate to different nucleosome repeat lengths.

### 6.6.1 Solenoid Chromatin Fiber Models

The classical solenoid fiber model has a one-start (1,1) helical organization of the chain, in which consecutive nucleosomes stack on top of each other (Figure 6.4a)
The interactions between nucleosomes adjacent on the DNA require bending of the intervening linker DNA. This is energetically unfavorable and could be facilitated by association with linker histones [8, 19]. Alternative conformations with straight linker DNA are discussed below. To allow for a higher nucleosome density than 6–7 nucleosomes per 11 nm fiber reported for the (1,1) conformation, other solenoid models were proposed [3, 9, 204]. These are characterized by an interdigitation of nucleosomes between adjacent turns of the helix, but differ in the nucleosome tilt angle with respect to the chromatin fiber axis [72]. In the fiber conformations proposed by Daban nucleosomes have high tilt angles of 40–60° forming (n,1) fibers with n=3–6 [204]. Examples for a (6,1) conformation of this type are shown in Figure 6.4e-g for different NRLs. In the model from the

Figure 6.4  Models of different chromatin fiber model conformations. (a) Classical solenoid model [31]. The nucleosomal DNA is colored in light violet cyan and the linker DNA in yellow (image from Ref. [68]). (b) Two-start helix with crossed-linker DNA (image from Ref. [68]). The color coding corresponds to that in (a). (c) Interdigitated solenoid model with low nucleosome tilt angle (image from Ref. [9]). Alternating nucleosome pairs are colored blue and magenta, and the nucleosomes at positions 1, 2, and 7 of the chain are indicated. (d) Two-start helix crossed-linker DNA fiber derived by extending the tetranucleosome crystal structure [65] (image from Ref. [9]). The color coding is the same as in (c). (e) Monte Carlo simulations of (6,1) fiber conformations with relatively high nucleosome tilt angles and different NRLs that are based on the conformation proposed by Daban [204]. The left structure shows the initial configurations and the right fiber is a representative conformation in thermal equilibrium obtained after MC simulations [72]. NRL = 189 bp, linear mass density ~ 7.6 nucleosomes/11 nm fiber, diameter ~ 33 nm. (f) Same as in (e) but for NRL = 199 bp; linear mass density ~ 6.9 nucleosomes/11 nm fiber, diameter ~ 36 nm. (g) Same as in (e) but for NRL = 207 bp. The initial fiber structure transformed into a random aggregate at thermal equilibrium. This is due to the increased electrostatic repulsion of the longer linker DNA.
Rhodes group the tilt angle is $\sim 20^\circ$ and the nucleosome stacking follows a zig-zag path, which cannot be described in terms of nucleosome stacks (Figure 6.4c) [3, 9]. The high compaction ratios of interdigitated fibers were experimentally observed in the electron microscopy study that identified two distinct structural classes of fibers [9]. For NRLs of 187–207 bp a diameter of 33–34 nm and a nucleosome packing ratio of $\sim 11$ nucleosomes per 11 nm fiber was measured. Longer repeat lengths of 217–237 bp associated into thicker fibers with a diameter of $\sim 44$ nm and a linear mass density of $\sim 15$ nucleosomes per 11 nm fiber.

6.6.2 Chromatin Fibers with Crossed Linker DNA

In crossed-linker DNA chromatin fibers nucleosomes interact with each other that are not adjacent on the nucleosome chain. This allows for straight linker DNA with crossings in the interior of the fiber along a zig-zag path as in the (7,3) and (7,4)
conformations shown in Figure 6.3b and the two-start fibers with straight linker DNA in Figure 6.4b, d. In these structures the fiber diameter would be expected to depend linearly on the length of the linker DNA [211]. However, results on this issue are contradictory. An increase of diameter with NRL was observed in two studies [13, 211], while others reported no change [29, 34] or an increase of the fiber diameter by \( \sim 10 \text{ nm} \) only between an NRL of 207 bp and 217 bp [9].

An experimentally well-established folding state is the crossed-linker DNA two-start chromatin fiber conformation with stacking of nucleosomes \( i \) and \( i+2 \) and adjacent nucleosomes connected by more or less straight linker DNA, that is, a (2,1) geometry [13, 65, 68, 212]. For this type of fiber, the nucleosome orientations and path of the linker DNA can be derived from the crystal structure of a tetranucleosome at 167 bp NRL and in the absence of linker histones [65]. The structure has a resolution of 9 \( \text{Å} \), and was solved by molecular replacement with the high-resolution nucleosome core structure determined previously [121]. The tetranucleosome structure can be extended into a continuous fiber [65], the stability of which was investigated by Monte Carlo simulations [72]. The resulting conformation at thermal equilibrium is in good agreement with structures observed by electron microscopy for NRLs of 167 and 197 bp [16].

The crossed-linker DNA fiber conformation is compatible with other fiber types that vary with respect to the number of nucleosome stacks and the DNA path. Conformations were proposed with (3,1), (5,2), and (7,3) geometry [208, 213, 214]. Thus, the potential conformational variability is significant. A central issue in this context is it to evaluate the energetics of the geometrically possible structures. One approach to systematically search for stable fiber conformations is illustrated in Figure 6.5 [209]. It shows an energy-minimized phase diagram based on the (2,1) fiber conformation derived for chicken erythrocyte chromatin fibers by Woodcock and coworkers [11, 203]. By varying the local nucleosome geometry a number of additional fiber structures were found to be stable in the computer simulations [209].

### 6.7 Conclusions

It is apparent from the studies reviewed here that the nucleosome chain is polymorphic and can organize into a variety of conformations. Its folding depends on a number of parameters including the nucleosome positioning (with respect to both its regularity and spacing), the protein composition (histone variants, presence/type of linker histone, other architectural proteins) as well as post-translational histone modifications. It is noteworthy that relatively small local variations can translate into large changes of the overall chain conformation. This is illustrated by the model for the reorganization of the nucleosome chain in response to a change of the local nucleosome geometry that could be for example brought about by binding of linker histone H1 (Figure 6.6). In its absence an open (2,1) fiber conformation is present for which a conformation derived from that of the tetranucleosome crystal structure with an NRL extended to \( \sim 189 \text{ bp} \) is depicted. In this...
state most of the linker DNA would be accessible for the binding of other protein factors. By constraining the local geometry of the DNA at the nucleosome entry–exit site, H1 could induce the compaction of the chain into a \( (6,1) \) fiber. In this conformation the linker DNA is located in the interior of the fiber so that association of other factors to this part would be inhibited. Thus, access to the linker DNA sequence can be regulated by the differential folding of the nucleosome chain. Its compaction can vary about tenfold from 1 to 2 nucleosomes per 11 nm in a very open conformation [14] up to 17 nucleosomes per 11 nm fiber in its fully compacted state [9]. This is in line with experimental studies of linker DNA access that showed large differences in protein binding between the folded and unfolded 17mer nucleosomal array [4]. To further evaluate the potentially large impact of the nucleosome chain conformation to regulate DNA access and associated molecular biological processes requires future studies that address a number of critical points:

1. The conformation(s) and effects of linker histone binding to the nucleosome and associated linker DNA remain to be further investigated.

Figure 6.6  Model for chromatin fiber compaction induced by binding of linker histone H1. All-atom model structures of a nucleosome with and without linker histone H1 (left panel) were used to build two corresponding coarse-grained models of a chain with 100 nucleosomes (right panel) that was subjected to Monte Carlo simulations. The change of the DNA geometry due to binding of linker histone H1 at the DNA entry–exit site of the nucleosome leads to a compaction of the chain into a condensed fiber structure with a diameter of about 30 nm [72].
2. The strength of the nucleosome–nucleosome interaction potential and its
dependence on the spatial orientation of interacting nucleosomes have to be
determined more precisely.
3. While experimentally challenging, it will be necessary to extend current *in vitro*
studies to larger chromatin fragments and higher nucleosome concentration.
4. The identification of the relevant cellular chromatin states in terms of protein
composition and histone modifications needs to be further advanced and
considered for the structural studies.

Advances in these four areas will be crucial to derive a more comprehensive
quantitative description of nucleosome chain folding to understand its organiza-
tion *in vitro* as well as in the cell nucleus.

**Acknowledgments**

I am grateful to Nick Kepper, Rene Stehr, Gero Wedemann, Ramona Ettig, and
Vladimir Teif for discussions. Work on chromatin conformation in my lab is
supported by the BMBF as a partner of the ERASysBio+ initiative supported
under the EU ERA-NET Plus scheme in FP7.

**References**

organization: balancing stability and plasticity. *Biochim Biophys Acta*, 1783,
2061–2079.
access to DNA target sites in folded chromatin fibers. *J Mol Biol*, 379,
772–786.
solution: determinants, mechanisms, and functions. *Annu Rev Biophys
Biomol Struct*, 31, 361–392.
chromatin structure and function: H1 stoichiometry and nucleosome repeat
measurements define the dimensions of the “30-nm” chromatin fiber:
evidence for a compact, interdigitated structure. *Proc Natl Acad Sci USA*, 103,
6506–6511.
structure studied by neutron scattering and scanning transmission electron
histone form a unique structural motif that directs the higher-order folding


51 Lantermann, A.B., Straub, T., Straffor, A., Yuan, G.C., Ekwall, K., and Korber,


Simpson, R.T. (1978) Structure of the chromatosome, a chromatin particle
containing 160 base pairs of DNA and all the histones. *Biochemistry*, 17, 5524–5531.


